

Convexity and Star-shapedness of Real Linear Images of Special Orthogonal Orbits

Pan-Shun Lau^{*1}, Tuen-Wai Ng^{†1}, and Nam-Kiu Tsing^{‡1}

¹Department of Mathematics, The University of Hong Kong, Pokfulam, Hong Kong

August 23, 2016

Abstract

Let $A \in \mathbb{R}^{n \times n}$ and $\text{SO}_n := \{U \in \mathbb{R}^{n \times n} : UU^t = I_n, \det U > 0\}$ be the set of $n \times n$ special orthogonal matrices. Define the (real) special orthogonal orbit of A by

$$O(A) := \{UAV : U, V \in \text{SO}_n\}.$$

In this paper, we show that the linear image of $O(A)$ is star-shaped with respect to the origin for arbitrary linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ if $n \geq 2^{\ell-1}$. In particular, for linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$ and when A has distinct singular values, we study $B \in O(A)$ such that $L(B)$ is a boundary point of $L(O(A))$. This gives an alternative proof of a result by Li and Tam on the convexity of $L(O(A))$ for linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$.

AMS Classification: 15A04, 15A18.

Keywords: linear transformation, special orthogonal orbits, convexity, star-shapedness

1 Introduction

Let $\mathcal{O}_n := \{U \in \mathbb{R}^{n \times n} : U^t U = U U^t = I_n\}$ and $\text{SO}_n := \{U \in \mathcal{O}_n : \det U > 0\}$ be the sets of $n \times n$ orthogonal matrices and $n \times n$ special orthogonal matrices respectively. For any $A \in \mathbb{R}^{n \times n}$, we define the special orthogonal orbit of A by

$$O(A) := \{UAV : U, V \in \text{SO}_n\}.$$

It is clear that every element in $O(A)$ has the same collection of singular values and the same sign of determinant. In [9], Thompson studied the set of diagonals of the matrices in $O(A)$, and in [8], Miranda and Thompson studied the

^{*}panlau@hku.hk

[†]ntw@maths.hku.hk

[‡]nktsing@hku.hk

characterizations of extreme values of $L(O(A))$ where $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ is a linear map.

A set S is said to be star-shaped with respect to $c \in S$ if for all $0 \leq \alpha \leq 1$ and $x \in S$, $\alpha x + (1-\alpha)c \in S$. The c is called a star center of S . In this paper, we shall study the star-shapedness of images of $O(A)$ under arbitrary linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$.

In fact the study of linear images of matrix orbits is a popular topic. If A, C are $n \times n$ complex matrices and \mathcal{U}_n denotes the group of $n \times n$ (complex) unitary matrices, then the (classical) numerical range of A , denoted by $W(A)$, and the C -numerical range of A , denoted by $W_C(A)$, are simply the images of the unitary orbit of A , denoted by

$$\mathcal{U}_n(A) := \{U^*AU : U \in \mathcal{U}_n\},$$

under the linear maps

$$X \mapsto \text{tr}(E_1 X) \quad \text{and} \quad X \mapsto \text{tr}(CX)$$

respectively, where E_1 is the diagonal matrix with diagonal entries $1, 0, \dots, 0$. It has been proved that $W(A)$ is always convex and $W_C(A)$ is always star-shaped (see [1], [2], [10]). Many results on the convexity and the star-shapedness of other generalized numerical ranges, which can be expressed as some particular linear images of $\mathcal{U}_n(A)$, have been obtained (e.g., see [1], [3], [4], [5], [6], [11], [12]).

Our paper is organized as follows. In Section 2, we study an inclusion relation of $L(O(A))$ with $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ and $n \geq 2^{\ell-1}$. We then apply the inclusion relation to show that $L(O(A))$ is star-shaped for general A and $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ where $n \geq 2^{\ell-1}$. In particular, the star-shapedness holds for $L(O(A))$ with $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$ and $n \geq 3$. Moreover, we shall extend our results to linear images of the following joint (real) orthogonal orbits,

$$\begin{aligned} \mathcal{O}_1(A_1, \dots, A_m; G) &:= \{(A_1 V, \dots, A_m V) : V \in G\}, \\ \mathcal{O}_2(A_1, \dots, A_m; G) &:= \{(U A_1, \dots, U A_m) : U \in G\}, \\ \mathcal{O}_3(A_1, \dots, A_m; G) &:= \{(U A_1 V, \dots, U A_m V) : U, V \in G\}, \end{aligned}$$

where $G = \mathcal{O}_n$ or SO_n . In Section 3, we study boundary points of $L(O(A))$ with $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$. When $A \in \mathbb{R}^{n \times n}$ has distinct singular values, we shall discuss the conditions on $U, V \in \text{SO}_n$ under which $L(UAV)$ will be a boundary point of $L(O(A))$. Then we show that the intersection of $L(O(A))$ and any of its supporting lines is path-connected. Combining the result in Section 2, convexity of $L(O(A))$ for $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$ then follows. This result was proved by Li and Tam [7] with a different approach. We shall also discuss the convexity of linear images of joint orthogonal orbits.

2 Star-shapedness of linear image of $O(A)$

The following is the first main theorem in this section.

Theorem 2.1. *Let $\ell \geq 3$. For any $A \in \mathbb{R}^{n \times n}$ and any linear map $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ with $n \geq 2^{\ell-1}$, $L(O(A))$ is star-shaped with respect to the origin.*

We need some lemmas to prove Theorem 2.1. Note that any linear map $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ can be expressed as

$$L(X) = (\text{tr}(P_1 X), \dots, \text{tr}(P_\ell X))^t$$

for some $P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$. For convenience, for $M \subseteq \mathbb{R}^{n \times n}$ and any $P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$, we define

$$\mathcal{L}(P_1, \dots, P_\ell; M) := \{(\text{tr}(P_1 X), \dots, \text{tr}(P_\ell X))^t : X \in M\}.$$

For $A, P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$, we let $\mathcal{S}_A(P_1, \dots, P_\ell)$ be the set containing (P'_1, \dots, P'_ℓ) where $P'_1, \dots, P'_\ell \in \mathbb{R}^{n \times n}$ and $\mathcal{L}(P'_1, \dots, P'_\ell; O(A)) \subseteq \mathcal{L}(P_1, \dots, P_\ell; O(A))$. This definition is motivated by Cheung and Tsing [1]. Below are some basic properties of $\mathcal{S}_A(P_1, \dots, P_\ell)$.

Lemma 2.2. *Let $A \in \mathbb{R}^{n \times n}$. For any $P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$, the followings hold:*

- (a) $\mathcal{S}_{XAY}(UP_1 V, \dots, UP_\ell V) = \mathcal{S}_A(P_1, \dots, P_\ell)$ for any $U, V, X, Y \in \text{SO}_n$;
- (b) $(UP_1 V, \dots, UP_\ell V) \in \mathcal{S}_A(P_1, \dots, P_\ell)$, for any $U, V \in \text{SO}_n$;
- (c) $\mathcal{S}_A(P'_1, \dots, P'_\ell) \subseteq \mathcal{S}_A(P_1, \dots, P_\ell)$ for any $(P'_1, \dots, P'_\ell) \in \mathcal{S}_A(P_1, \dots, P_\ell)$;
- (d) $\mathcal{L}(P_1, \dots, P_\ell; O(A)) = \{(\text{tr}(P'_1 A), \dots, \text{tr}(P'_\ell A))^t : (P'_1, \dots, P'_\ell) \in \mathcal{S}_A(P_1, \dots, P_\ell)\}.$

Proof. (a), (b) and (c) are trivial. For (d), “ \subseteq ” follows from (b) and “ \supseteq ” follows from the definition of $\mathcal{S}_A(P_1, \dots, P_\ell)$. \square

Lemma 2.3. *The following statements are equivalent (hence if one of these statements holds then the other three must also hold):*

- (a) $L(O(A))$ is star-shaped with respect to the origin for any $A \in \mathbb{R}^{n \times n}$ and any linear map $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$;
- (b) $\mathcal{S}_A(P_1, \dots, P_\ell)$ is star-shaped with respect to $(0_n, \dots, 0_n)$ for any $A \in \mathbb{R}^{n \times n}$ and any $P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$, where 0_n is the $n \times n$ zero matrix;
- (c) $L(\text{SO}_n)$ is star-shaped with respect to the origin for any linear map $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$;
- (d) $\mathcal{S}_{I_n}(P_1, \dots, P_\ell)$ is star-shaped with respect to $(0_n, \dots, 0_n)$ for any $P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$.

Proof. ((a) \Rightarrow (b)) For any $(P'_1, \dots, P'_\ell) \in \mathcal{S}_A(P_1, \dots, P_\ell)$, $U, V \in \text{SO}_n$ and $0 \leq \alpha \leq 1$, we have

$$(\text{tr}(\alpha P'_1 U A V), \dots, \text{tr}(\alpha P'_\ell U A V))^t \in \mathcal{L}(P'_1, \dots, P'_\ell; O(A)) \subseteq \mathcal{L}(P_1, \dots, P_\ell; O(A)).$$

Hence $\alpha(P'_1, \dots, P'_\ell) \in \mathcal{S}_A(P_1, \dots, P_\ell)$.

((b) \Rightarrow (a)) Apply Lemma 2.2 (b).

((a) \Rightarrow (c)) If we take $A = I_n$, then $O(A) = \text{SO}_n$.

((c) \Rightarrow (a)) Let $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ be linear and $A \in \mathbb{R}^{n \times n}$. For any $U \in \text{SO}_n$, define linear map $L_{UA} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ by

$$L_{UA}(X) = L(UAX).$$

For any $U, V \in \text{SO}_n$ and $0 \leq \alpha \leq 1$, since $L_{UA}(\text{SO}_n)$ is star-shaped with respect to the origin, there exists $V' \in \text{SO}_n$ such that

$$\alpha L(UAV) = \alpha L_{UA}(V) = L_{UA}(V') = L(UAV') \in L(O(A)).$$

((c) \Leftrightarrow (d)) Apply similar arguments as those in (a) \Leftrightarrow (b). \square

To prove Theorem 2.1, we apply Lemma 2.3 and show the star-shapedness of $\mathcal{S}_{I_n}(P_1, \dots, P_\ell)$ for any $P_1, \dots, P_\ell \in \mathbb{R}^{n \times n}$ with $n \geq 2^{\ell-1}$. For simplicity, we denote $\mathcal{S}_{I_n}(P_1, \dots, P_\ell)$ by $\mathcal{S}(P_1, \dots, P_\ell)$. In fact, by the following lemma, we may focus only on the case of $n = 2^{\ell-1}$.

Lemma 2.4. *If $\mathcal{S}(\hat{P}_1, \dots, \hat{P}_\ell)$ is star-shaped with respect to the origin for all $\hat{P}_1, \dots, \hat{P}_\ell \in \mathbb{R}^{n \times n}$, then for all $m > n$ and for all $P_1, \dots, P_\ell \in \mathbb{R}^{m \times m}$, $\mathcal{S}(P_1, \dots, P_\ell)$ is star-shaped with respect to the origin.*

Proof. Let $m = n + k$ where k is a positive integer. For any $(P'_1, \dots, P'_\ell) \in \mathcal{S}(P_1, \dots, P_\ell)$, we write

$$P'_i = \begin{bmatrix} P'_{i1} & P'_{i2} \\ P'_{i3} & P'_{i4} \end{bmatrix},$$

where $P'_{i1} \in \mathbb{R}^{n \times n}$ and $P'_{i4} \in \mathbb{R}^{k \times k}$. We shall show that $(P'_1(\epsilon), \dots, P'_\ell(\epsilon)) \in \mathcal{S}(P_1, \dots, P_\ell)$ where $P'_i(\epsilon) = (\epsilon I_n \oplus I_k)P'_i$ and $0 \leq \epsilon \leq 1$. For any $U \in \text{SO}_m$, we write

$$U = \begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix},$$

where $U_1 \in \mathbb{R}^{n \times n}$ and $U_4 \in \mathbb{R}^{k \times k}$. Then for $0 \leq \epsilon \leq 1$, by the hypothesis of the lemma, there exists $V \in \text{SO}_n$ such that

$$\begin{aligned} & (\text{tr}(P'_1(\epsilon)U), \dots, \text{tr}(P'_\ell(\epsilon)U))^t \\ &= \epsilon (\text{tr}(P'_{11}U_1 + P'_{12}U_3), \dots, \text{tr}(P'_{\ell 1}U_1 + P'_{\ell 2}U_3))^t \\ & \quad + (\text{tr}(P'_{13}U_2 + P'_{14}U_4), \dots, \text{tr}(P'_{\ell 3}U_2 + P'_{\ell 4}U_4))^t \\ &= \left(\text{tr}[(P'_{11}U_1 + P'_{12}U_3)V], \dots, \text{tr}[(P'_{\ell 1}U_1 + P'_{\ell 2}U_3)V] \right)^t \\ & \quad + \left(\text{tr}(P'_{13}U_2 + P'_{14}U_4), \dots, \text{tr}(P'_{\ell 3}U_2 + P'_{\ell 4}U_4) \right)^t \\ &= \left(\text{tr}[P'_1 U(V \oplus I_k)], \dots, \text{tr}[P'_\ell U(V \oplus I_k)] \right)^t \\ &\in \mathcal{L}(P'_1, \dots, P'_\ell; \text{SO}_m) \\ &\subseteq \mathcal{L}(P_1, \dots, P_\ell; \text{SO}_m). \end{aligned}$$

Since this holds for all $U \in \text{SO}_m$, we have $(P'_1(\epsilon), \dots, P'_\ell(\epsilon)) \in \mathcal{S}(P_1, \dots, P_\ell)$. Note that the preceding result also holds if we multiply arbitrary n rows of P'_i by $0 \leq \epsilon \leq 1$. We re-apply the result by considering all n -combinations of rows to obtain $\epsilon^N(P'_1, \dots, P'_\ell) \in \mathcal{S}(P_1, \dots, P_\ell)$, where $N = \frac{m!}{n!k!}$. For any $0 \leq \alpha \leq 1$, we put $\epsilon = \sqrt[n]{\alpha}$ to obtain $\alpha(P'_1, \dots, P'_\ell) \in \mathcal{S}(P_1, \dots, P_\ell)$. \square

We now consider the following recursively defined matrices. Let

$$R(\theta_1) = \begin{bmatrix} \cos \theta_1 & \sin \theta_1 \\ -\sin \theta_1 & \cos \theta_1 \end{bmatrix}$$

and

$$R(\theta_1, \dots, \theta_k) = \begin{bmatrix} \cos \theta_k I_N & \sin \theta_k R(\theta_1, \dots, \theta_{k-1}) \\ -\sin \theta_k R(\theta_1, \dots, \theta_{k-1})^t & \cos \theta_k I_N \end{bmatrix}$$

where $N = 2^{k-1}$. Note that $R(\theta_1, \dots, \theta_k) \in \text{SO}_{2^k}$.

Lemma 2.5. *Let $\ell \geq 2$ and $P_1, \dots, P_\ell \in \mathbb{R}^{N \times N}$ where $N = 2^{\ell-1}$. Then for any $U, V \in \text{SO}_N$, the set*

$$E(U, V) :=$$

$$\left\{ \left(\text{tr}(R(\theta_1, \dots, \theta_{\ell-1})UP_1V), \dots, \text{tr}(R(\theta_1, \dots, \theta_{\ell-1})UP_\ell V) \right)^t : \theta_1, \dots, \theta_{\ell-1} \in [0, 2\pi] \right\}$$

is an ellipsoid in \mathbb{R}^ℓ centered at the origin and is a subset of $\mathcal{L}(P_1, \dots, P_\ell; \text{SO}_N)$.

Proof. We first show that for any $A \in \mathbb{R}^{N \times N}$ where $N = 2^{\ell-1}$,

$$\text{tr}(R(\theta_1, \dots, \theta_{\ell-1})A) = \begin{bmatrix} a_1 & \dots & a_\ell \end{bmatrix} \begin{bmatrix} \cos \theta_{\ell-1} \\ \sin \theta_{\ell-1} \cos \theta_{\ell-2} \\ \sin \theta_{\ell-1} \sin \theta_{\ell-2} \cos \theta_{\ell-3} \\ \vdots \\ \sin \theta_{\ell-1} \sin \theta_{\ell-2} \dots \sin \theta_1 \end{bmatrix}$$

for some $a_1, \dots, a_\ell \in \mathbb{R}$ by induction on ℓ . The case for $\ell = 2$ is trivial. Now assume it is true for $\ell \leq m$ where $m \geq 2$ and consider $A \in \mathbb{R}^{2^M \times 2^M}$ where $M = 2^{m-1}$. We write

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

where $A_i \in \mathbb{R}^{M \times M}$, $i = 1, \dots, 4$. Then

$$\text{tr}(R(\theta_1, \dots, \theta_m)A) = \cos \theta_m \text{tr}(A_1 + A_4) + \sin \theta_m \text{tr}(R(\theta_1, \dots, \theta_{m-1})(A_3 - A_2^t)).$$

By induction assumption on $\text{tr}(R(\theta_1, \dots, \theta_{m-1})(A_3 - A_2^t))$, $\text{tr}(R(\theta_1, \dots, \theta_m)A)$ is in the desired form. Hence we have

$$E(U, V) = \left\{ T \begin{bmatrix} \cos \theta_{\ell-1} \\ \sin \theta_{\ell-1} \cos \theta_{\ell-2} \\ \sin \theta_{\ell-1} \sin \theta_{\ell-2} \cos \theta_{\ell-3} \\ \vdots \\ \sin \theta_{\ell-1} \sin \theta_{\ell-2} \dots \sin \theta_1 \end{bmatrix} : \theta_1, \dots, \theta_{\ell-1} \in [0, 2\pi] \right\},$$

for some $T \in \mathbb{R}^{\ell \times \ell}$ and hence $E(U, V)$ is an ellipsoid in \mathbb{R}^ℓ centered at the origin. As $R(\theta_1, \dots, \theta_k)$ is a special orthogonal matrix, $E(U, V) \subseteq \mathcal{L}(P_1, \dots, P_\ell; \text{SO}_N)$. \square

Lemma 2.6. *Let $\ell \geq 3$. For any $P_1, \dots, P_\ell \in \mathbb{R}^{N \times N}$ where $N = 2^{\ell-1}$, there exist $U, V \in \text{SO}_N$ such that $E(U, V)$ defined in Lemma 2.5 degenerates (i.e., $E(U, V)$ is contained in an affine hyperplane in \mathbb{R}^ℓ).*

Proof. From the proof of Lemma 2.5, we see that if there exist $U, V \in \text{SO}_N$ such that

$$UP_1V = \begin{bmatrix} P_1^{(1)} & P_2^{(1)} \\ P_3^{(1)} & P_4^{(1)} \end{bmatrix}$$

where $P_i^{(1)} \in \mathbb{R}^{\frac{N}{2} \times \frac{N}{2}}$, $i = 1, \dots, 4$, $\text{tr}(P_1^{(1)} + P_4^{(1)}) = 0$ and $P_2^{(1)} = P_3^{(1)} = 0$, then the first coordinate of $E(U, V)$ is always 0 and hence $E(U, V)$ degenerates. Let $U', V' \in \text{SO}_N$ be such that $U'P_1V' = \text{diag}(p_1, \dots, p_N)$. Then

$$U = U', \quad V = V' \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \oplus \dots \oplus \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right)$$

will give the desired UP_1V . \square

Note that, by considering $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $P_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, then for any $U, V \in \text{SO}_2$, the ellipse $E(U, V)$ defined in Lemma 2.5 is always non-degenerate. Hence Lemma 2.6 and Theorem 2.1 fail to hold for $\ell = 2$.

We are now ready to prove our first main result.

Proof of Theorem 2.1. By Lemma 2.3 and Lemma 2.4, it suffices to show that for any $P_1, \dots, P_\ell \in \mathbb{R}^{N \times N}$ with $N = 2^{\ell-1}$, $\mathcal{S}(P_1, \dots, P_\ell)$ is star-shaped with respect to $(0_N, \dots, 0_N)$. Let $(P'_1, \dots, P'_\ell) \in \mathcal{S}(P_1, \dots, P_\ell)$ and $0 \leq \alpha \leq 1$. For any $U \in \text{SO}_N$, we define $E(I_N, U)$ as in Lemma 2.5. If $\alpha(\text{tr}(P'_1U), \dots, \text{tr}(P'_\ell U))^t \in E(I_N, U)$, then we have

$$\alpha(\text{tr}(P'_1U), \dots, \text{tr}(P'_\ell U))^t \in \mathcal{L}(P'_1, \dots, P'_\ell; \text{SO}_N) \subseteq \mathcal{L}(P_1, \dots, P_\ell; \text{SO}_N).$$

Assume now $\alpha(\text{tr}(P'_1U), \dots, \text{tr}(P'_\ell U))^t \notin E(I_N, U)$. As the center of $E(I_N, U)$ is the origin, we have $\alpha(\text{tr}(P'_1U), \dots, \text{tr}(P'_\ell U))^t$ lies inside the ellipsoid $E(I_N, U)$. As $\text{SO}_N \times \text{SO}_N$ is path connected, consider a continuous function $f : [0, 1] \rightarrow \text{SO}_N \times \text{SO}_N$ with $f(0) = (I_N, U)$ and $f(1) = (U', V')$ where (U', V') are defined in Lemma 2.6. Then by continuity of f , there exists $s \in [0, 1]$ such that $\alpha(\text{tr}(P'_1U), \dots, \text{tr}(P'_\ell U))^t \in E(f(s)) \subseteq \mathcal{L}(P'_1, \dots, P'_\ell; \text{SO}_N) \subseteq \mathcal{L}(P_1, \dots, P_\ell; \text{SO}_N)$. As it is true for all $U \in \text{SO}_N$, we have

$$\alpha(P'_1, \dots, P'_\ell) + (1 - \alpha)(0_n, \dots, 0_n) = \alpha(P'_1, \dots, P'_\ell) \in \mathcal{S}(P_1, \dots, P_\ell).$$

\square

In fact for $\ell = 2$, we have the following theorem, the proof of which is given by Lemma 2.8 to Corollary 2.11.

Theorem 2.7. *Let $A \in \mathbb{R}^{n \times n}$ and $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$ be a linear map with $n \geq 3$. Then $L(O(A))$ is star-shaped with respect to the origin.*

Lemma 2.8. *Let $n \geq 2$. For any $P, Q \in \mathbb{R}^{n \times n}$, $U \in \text{SO}_n$, the locus of the point $(\text{tr}(T_\theta P U), \text{tr}(T_\theta Q U))^t$ where $T_\theta = R(\theta) \oplus I_{n-2}$ forms an ellipse $E(U)$ in \mathbb{R}^2 when θ runs through $[0, 2\pi]$.*

Proof. We write

$$P = \begin{bmatrix} \frac{p_{(1)}}{P_{(3)}} \\ \frac{p_{(2)}}{P_{(3)}} \\ \frac{p_{(3)}}{P_{(3)}} \end{bmatrix}, \quad Q = \begin{bmatrix} \frac{q_{(1)}}{Q_{(3)}} \\ \frac{q_{(2)}}{Q_{(3)}} \\ \frac{q_{(3)}}{Q_{(3)}} \end{bmatrix} \quad \text{and} \quad U = [\begin{array}{c|c|c} u^{(1)} & u^{(2)} & U^{(3)} \end{array}]$$

where $p_{(1)}^t, p_{(2)}^t, q_{(1)}^t, q_{(2)}^t, u^{(1)}, u^{(2)} \in \mathbb{R}^n$ and $P_{(3)}^t, Q_{(3)}^t, U^{(3)} \in \mathbb{R}^{n \times (n-2)}$. Direct computation shows

$$\text{tr}(T_\theta P U) = \cos \theta (p_{(1)} u^{(1)} + p_{(2)} u^{(2)}) + \sin \theta (p_{(2)} u^{(1)} - p_{(1)} u^{(2)}) + \text{tr}(P_{(3)}^t U^{(3)}).$$

Similarly for $\text{tr}(T_\theta Q U)$. Hence

$$\begin{bmatrix} \text{tr}(T_\theta P U) \\ \text{tr}(T_\theta Q U) \end{bmatrix} = \begin{bmatrix} p_{(1)} u^{(1)} + p_{(2)} u^{(2)} & p_{(2)} u^{(1)} - p_{(1)} u^{(2)} \\ q_{(1)} u^{(1)} + q_{(2)} u^{(2)} & q_{(2)} u^{(1)} - q_{(1)} u^{(2)} \end{bmatrix} \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} + \begin{bmatrix} \text{tr}(P_{(3)}^t U^{(3)}) \\ \text{tr}(Q_{(3)}^t U^{(3)}) \end{bmatrix},$$

the locus of which forms an ellipse (possibly degenerate) when θ runs through $[0, 2\pi]$. \square

Lemma 2.9. *For any $P, Q \in \mathbb{R}^{n \times n}$ with $n \geq 3$, there exists $U_0 \in \text{SO}_n$ such that the ellipse $E(U_0)$ defined in Lemma 2.8 degenerates.*

Proof. Note that $E(U)$ degenerates if we find orthonormal vectors $u^{(1)}, u^{(2)} \in \mathbb{R}^n$ such that the matrix

$$\begin{bmatrix} p_{(1)} u^{(1)} + p_{(2)} u^{(2)} & p_{(2)} u^{(1)} - p_{(1)} u^{(2)} \\ q_{(1)} u^{(1)} + q_{(2)} u^{(2)} & q_{(2)} u^{(1)} - q_{(1)} u^{(2)} \end{bmatrix}$$

is singular. We will show that for any given $p_1, p_2 \in \mathbb{R}^n$, there exist orthonormal vectors u_1, u_2 such that $p_1^t u_2 = p_2^t u_1 = p_1^t u_1 + p_2^t u_2 = 0$. By scaling and rotating, we assume without loss of generality that $p_1 = (1, 0, \dots, 0)^t$ and $p_2 = (a, b, 0, \dots, 0)^t$ where $a, b \in \mathbb{R}$ and $0 \leq b \leq 1$. If $a = 0$ or $b = 0$, we can take $u_1 = (-b, 0, \sqrt{1-b^2}, 0, \dots, 0)^t$ and $u_2 = (0, 1, 0, \dots, 0)^t$. Now, assume that $a \neq 0$ and $0 < b \leq 1$. For $\theta \in [0, \pi]$ consider unit vectors

$$v_\theta = \begin{bmatrix} 0 \\ \cos \theta \\ \sin \theta \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad w_\theta = \frac{1}{\sqrt{b^2 \sin^2 \theta + a^2}} \begin{bmatrix} -b \sin \theta \\ a \sin \theta \\ -a \cos \theta \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Clearly, $p_1^t v_\theta = p_2^t w_\theta = v_\theta^t w_\theta = 0$. Define $f(\theta) = p_1 w_\theta + p_2 v_\theta = b \cos \theta - \frac{b \sin \theta}{\sqrt{b^2 \sin^2 \theta + a^2}}$ which is a continuous function with $f(0) = b$ and $f(\pi) = -b$. Hence there exists $\theta' \in [0, \pi]$ such that $f(\theta') = 0$. Then we take $u_2 = v_{\theta'}$ and $u_1 = w_{\theta'}$. \square

Lemma 2.10. For $P, Q \in \mathbb{R}^{n \times n}$, $n \geq 3$ and $0 \leq \epsilon \leq 1$ we define

$$P_\epsilon = \begin{bmatrix} \epsilon I_2 & \\ & I_{n-2} \end{bmatrix} P \quad \text{and} \quad Q_\epsilon = \begin{bmatrix} \epsilon I_2 & \\ & I_{n-2} \end{bmatrix} Q.$$

Then $(P_\epsilon, Q_\epsilon) \in \mathcal{S}(P, Q)$.

Proof. For any $U \in \text{SO}_n$, consider the ellipse $E(U)$ defined in Lemma 2.8. If $(\text{tr}(P_\epsilon U), \text{tr}(Q_\epsilon U))^t \in E(U)$, then we have $(\text{tr}(P_\epsilon U), \text{tr}(Q_\epsilon U))^t \in \mathcal{L}(P, Q; \text{SO}_n)$. Now assume that $(\text{tr}(P_\epsilon U), \text{tr}(Q_\epsilon U))^t \notin E(U)$. Then $(\text{tr}(P_\epsilon U), \text{tr}(Q_\epsilon U))^t$ lies inside the ellipse $E(U)$. Since SO_n is path-connected, consider a continuous function $f : [0, 1] \rightarrow \text{SO}_n$ with $f(0) = U$ and $f(1) = U_0$ where U_0 is defined in Lemma 2.9. Since $E(f(1))$ degenerates, by continuity of f , there exist $s \in [0, 1]$ such that $(\text{tr}(P_\epsilon U), \text{tr}(Q_\epsilon U))^t \in E(f(s)) \subseteq \mathcal{L}(P, Q; \text{SO}_n)$. As it is true for all $U \in \text{SO}_n$, we have $\mathcal{L}(P_\epsilon, Q_\epsilon; \text{SO}_n) \subseteq \mathcal{L}(P, Q; \text{SO}_n)$ and hence $(P_\epsilon, Q_\epsilon) \in \mathcal{S}(P, Q)$. \square

Lemma 2.10 remains valid if we consider $\mathcal{S}_A(P, Q)$ instead of $\mathcal{S}(P, Q)$.

Corollary 2.11. Let $A \in \mathbb{R}^{n \times n}$ and $n \geq 3$. For any $P, Q \in \mathbb{R}^{n \times n}$ and $0 \leq \epsilon \leq 1$, we define

$$P_\epsilon = \begin{bmatrix} \epsilon I_2 & \\ & I_{n-2} \end{bmatrix} P \quad \text{and} \quad Q_\epsilon = \begin{bmatrix} \epsilon I_2 & \\ & I_{n-2} \end{bmatrix} Q.$$

Then $(P_\epsilon, Q_\epsilon) \in \mathcal{S}_A(P, Q)$.

Proof. For any $U, V \in \text{SO}_n$, let $P' = P U A V$, $Q = Q U A V$, $P'_\epsilon = (\epsilon I_2 \oplus I_{n-2}) P' = P_\epsilon U A V$ and $Q'_\epsilon = (\epsilon I_2 \oplus I_{n-2}) Q' = Q_\epsilon U A V$. By Lemma 2.10, because $(P'_\epsilon, Q'_\epsilon) \in \mathcal{S}(P', Q')$, there exists $W \in \text{SO}_n$ such that

$$\begin{aligned} (\text{tr}(P_\epsilon U A V), \text{tr}(Q_\epsilon U A V))^t &= (\text{tr} P'_\epsilon, \text{tr} Q'_\epsilon)^t \\ &= (\text{tr}(P' W), \text{tr}(Q' W))^t \\ &= (\text{tr}(P U A V W), \text{tr}(Q U A V W))^t \\ &\in \mathcal{L}(P, Q; O(A)). \end{aligned}$$

As this is true for all $U, V \in \text{SO}_n$, we have $\mathcal{L}(P_\epsilon, Q_\epsilon; O(A)) \subseteq \mathcal{L}(P, Q; O(A))$. \square

Note that in Lemma 2.10 and Corollary 2.11, P_ϵ, Q_ϵ can be defined by picking arbitrary two rows of P and Q instead of the first two rows. We are now ready to prove our second main theorem.

Proof of Theorem 2.7. By Lemma 2.3, it suffices to show that for all $P, Q \in \mathbb{R}^{n \times n}$, $\mathcal{S}(P, Q)$ is star-shaped with respect to $(0_n, 0_n)$. Let $(P', Q') \in \mathcal{S}(P, Q)$ and $0 \leq \alpha \leq 1$. We apply Lemma 2.10 repeatedly to every two rows of P, Q . Then we have $(\epsilon^N P', \epsilon^N Q') \in \mathcal{S}(P', Q') \subseteq \mathcal{S}(P, Q)$ where $N = \frac{n!}{2(n-2)!}$. Taking $\epsilon = \sqrt[n]{\alpha}$, we have

$$\alpha(P', Q') = \alpha(P', Q') + (1 - \alpha)(0_n, 0_n) \in \mathcal{S}(P, Q).$$

□

For the case of $\ell = 2$ and $\ell = 3$, we know that $n = 3$ and $n = 4$ are respectively the smallest integers such that $L(O(A))$ is star-shaped for all $A \in \mathbb{R}^{n \times n}$ and all linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$. However, for $\ell \geq 4$, $n = 2^{\ell-1}$ may not be the smallest integer to ensure star-shapedness of $L(O(A))$. One may ask the following question.

Problem 1. For a given $\ell \geq 4$, what is the smallest n such that $L(\text{SO}_n)$ is star-shaped for all linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$?

The preceding results on star-shapedness of $L(O(A))$ can be easily generalized to the following joint orbits. We let $(\mathbb{R}^{n \times n})^m := \{(A_1, \dots, A_m) : A_1, \dots, A_m \in \mathbb{R}^{n \times n}\}$.

Definition 1. For any $A_1, \dots, A_m \in \mathbb{R}^{n \times n}$, we define

$$\begin{aligned} \mathcal{O}_1(A_1, \dots, A_m; G) &:= \{(A_1 V, \dots, A_m V) : V \in G\}, \\ \mathcal{O}_2(A_1, \dots, A_m; G) &:= \{(U A_1, \dots, U A_m) : U \in G\}, \\ \mathcal{O}_3(A_1, \dots, A_m; G) &:= \{(U A_1 V, \dots, U A_m V) : U, V \in G\}, \end{aligned}$$

where $G = \mathcal{O}_n$ or SO_n .

Theorem 2.12. Let $L : (\mathbb{R}^{n \times n})^m \rightarrow \mathbb{R}^\ell$ be linear, $(A_1, \dots, A_m) \in (\mathbb{R}^{n \times n})^m$ and $G = \mathcal{O}_n$ or SO_n . If

- (i) $\ell = 2$ and $n \geq 3$, or
- (ii) $\ell \geq 3$ and $n \geq 2^{\ell-1}$,

then $L(\mathcal{O}_i(A_1, \dots, A_m; G))$, $i = 1, 2, 3$, are star-shaped with respect to the origin.

Proof. The case of $G = \mathcal{O}_n$ can be derived from the case $G = \text{SO}_n$ easily. Hence we consider the case $G = \text{SO}_n$ only and simply denote $\mathcal{O}_i(A_1, \dots, A_m; \text{SO}_n)$ by $\mathcal{O}_i(A_1, \dots, A_m)$. For any given $L : (\mathbb{R}^{n \times n})^m \rightarrow \mathbb{R}^\ell$, express it by

$$L(X_1, \dots, X_m) = \left(\text{tr} \left(\sum_{i=1}^m P_i^{(1)} X_i \right), \dots, \text{tr} \left(\sum_{i=1}^m P_i^{(\ell)} X_i \right) \right)^t,$$

for some $P_i^{(j)} \in \mathbb{R}^{n \times n}$, $i = 1, \dots, m, j = 1, \dots, \ell$. For $\mathbf{O}_1(A_1, \dots, A_m)$ we have

$$\begin{aligned} & L(\mathbf{O}_1(A_1, \dots, A_m)) \\ &= \left\{ \left(\text{tr} \left(\sum_{i=1}^m P_i^{(1)} A_i U \right), \dots, \text{tr} \left(\sum_{i=1}^m P_i^{(\ell)} A_i U \right) \right)^t : U \in \text{SO}_n \right\} \\ &= \mathcal{L} \left(\sum_{i=1}^m P_i^{(1)} A_i, \dots, \sum_{i=1}^m P_i^{(\ell)} A_i; \text{SO}_n \right). \end{aligned}$$

Similarly for $L(\mathbf{O}_2(A_1, \dots, A_m))$. Hence the star-shapedness follows from Theorem 2.1 and Theorem 2.7.

Now consider the case of $\mathbf{O}_3(A_1, \dots, A_m)$. For any $U, V \in \text{SO}_n$, we have

$$\begin{aligned} L(UA_1V, \dots, UA_mV) &= \left(\text{tr} \left(\sum_{i=1}^m P_i^{(1)} UA_iV \right), \dots, \text{tr} \left(\sum_{i=1}^m P_i^{(\ell)} UA_iV \right) \right)^t \\ &\in \mathcal{L} \left(\sum_{i=1}^m P_i^{(1)} UA_i, \dots, \sum_{i=1}^m P_i^{(\ell)} UA_i; \text{SO}_N \right). \end{aligned}$$

By star-shapedness of $\mathcal{L} \left(\sum_{i=1}^m P_i^{(1)} UA_i, \dots, \sum_{i=1}^m P_i^{(\ell)} UA_i; \text{SO}_N \right)$, for any $0 \leq \alpha \leq 1$ we have

$$\begin{aligned} \alpha L(UA_1V, \dots, UA_mV) &\in \mathcal{L} \left(\sum_{i=1}^m P_i^{(1)} UA_i, \dots, \sum_{i=1}^m P_i^{(1)} UA_i; \text{SO}_N \right)^t \\ &\subseteq L(\mathbf{O}_3(A_1, \dots, A_m)). \end{aligned}$$

□

3 Convexity of linear image of $O(A)$

We first give two non-convex examples, one is a linear image of $O(A)$ under $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ with $\ell \geq 3$ and another is a linear image of $\mathbf{O}_3(A_1, \dots, A_m)$ under $L : (\mathbb{R}^{n \times n})^m \rightarrow \mathbb{R}^\ell$ with $\ell \geq 2$.

Example 1. Consider $O(I_n) = \text{SO}_n$ with $n \geq 2$ and the linear map $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ with $\ell \geq 3$ defined by

$$L(X) = (\text{tr}(P_1 X), \dots, \text{tr}(P_\ell X))^t$$

where

$$P_1 = I_{n-2} \oplus 0_2, \quad P_2 = I_{n-2} \oplus \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad P_3 = I_{n-2} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and $P_j = 0_n$ for $j = 4, \dots, \ell$. The mid-point of points $L(I_n) = (n-2, n-1, n-2, 0, \dots, 0)^t$ and $L\left(I_{n-2} \oplus \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right) = (n-2, n-2, n-1, 0, \dots, 0)^t$ is in $L(P_1, \dots, P_\ell; \text{SO}_n)$ only if there exists $U \in \text{SO}_n$ having the form

$$U = I_{n-2} \oplus \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}$$

with $u_{11} = \frac{1}{2} = u_{21}$. This is impossible as $u_{11}^2 + u_{21}^2 = 1$. Hence $L(\text{SO}_n)$ is non-convex.

Example 2. For $n \geq 3$, $m \geq 2$, $\ell \geq 2$, consider the matrices,

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \oplus 0_{n-3}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \oplus 0_{n-3}, \quad A_j = 0_n, \quad j = 3, \dots, m,$$

and the linear map $L : (\mathbb{R}^{n \times n})^m \rightarrow \mathbb{R}^\ell$ defined by

$$L(X_1, \dots, X_m) := (\text{tr}(A_1 X_1 + A_2 X_2), \text{tr}(A_2 X_1 - A_1 X_2), 0, \dots, 0)^t.$$

By taking $U = V = I_n$, and $U = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \oplus I_{n-3}$, $V = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \oplus I_{n-3}$

respectively, we have $(2, 0, 0, \dots, 0)^t, (0, 2, 0, \dots, 0)^t \in L(\mathcal{O}_3(A_1, \dots, A_m))$. We shall show that their mid-point which is $(1, 1, 0, \dots, 0)^t \notin L(\mathcal{O}_3(A_1, \dots, A_m))$. For any $U = [u_{ij}]$, $V = [v_{ij}] \in \text{SO}_n$, by direct computation we have

$$UA_1V = \begin{bmatrix} u_{11}v_{11} & * & * \\ * & u_{21}v_{12} & * \\ * & * & * \end{bmatrix}, \quad UA_2V = \begin{bmatrix} u_{12}v_{13} & * & * \\ * & u_{22}v_{22} & * \\ * & * & * \end{bmatrix}.$$

Hence $(1, 1, 0, \dots, 0) \in L(\mathcal{O}_3(A_1, \dots, A_m))$ only if $u_{11}v_{11} + u_{22}v_{22} = 1 = u_{21}v_{12} - u_{12}v_{13}$ for some $U, V \in \text{SO}_n$. We shall show that such U, V do not exist. For $X = (x_{ij})$, $Y = (y_{ij}) \in \mathbb{R}^{n \times n}$, denote $X \circ Y := (x_{ij}y_{ij}) \in \mathbb{R}^{n \times n}$. Since each absolute row (column) sum of $U \circ V$ is not greater than one, we have $(1, 1, 0, \dots, 0) \in L(\mathcal{O}_3(A_1, \dots, A_m))$ only if there exist $U, V \in \text{SO}_n$ such that

$$U \circ V = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & * \end{bmatrix} \quad \text{or} \quad U \circ V = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & * \end{bmatrix}.$$

The possible choices of the leading 2×2 principal submatrices of U and V are

$$\pm \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & k_1 \\ -k_2 & k_1 k_2 \end{bmatrix}$$

where $k_1, k_2 = \pm 1$. However, any two of them will not give the $U \circ V$ as required.

From the above two examples we know that $L(O(A))$ is not convex in general. However if the codomain of L is \mathbb{R}^2 then $L(O(A))$ is always convex. This result was obtained by Li and Tam [7] by using techniques in Lie algebra. In the following, we shall give an alternative proof on this result by showing that $L(O(A))$ has convex boundary for all $A \in \mathbb{R}^{n \times n}$ and linear $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$, i.e., the intersection of $L(O(A))$ with any of its supporting lines is path connected. Combining with the star-shapedness property of $L(O(A))$, the convexity of $L(O(A))$ follows. We first need some notations.

Definition 2. For $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, we denote its diagonal as $d(A) = (a_{11}, a_{22}, \dots, a_{nn})^t \in \mathbb{R}^n$. We further denote the sum of the first k diagonal elements of A by $t_k(A)$. Moreover for $P \in \mathbb{R}^{n \times n}$, we denote $r(P, A) = \max\{\text{tr}(PUAV) : U, V \in \text{SO}_n\}$ and $\mathcal{G}_P(A) = \{B \in O(A) : \text{tr}(PB) = r(P, A)\}$.

We shall characterize the set $\mathcal{G}_P(A)$ when A has distinct singular values and then show that it is path connected. Note that for any $U, V \in \text{SO}_n$, $\mathcal{G}_P(UAV) = \mathcal{G}_P(A)$ and $\mathcal{G}_{UPV}(A) = \{V^t B U^t : B \in \mathcal{G}_P(A)\}$. Hence we may assume that A, P are diagonal matrices.

Lemma 3.1. Let $A = \text{diag}(a_1, \dots, a_{n-1}, a_n)$ where $a_1 > a_2 > \dots > a_{n-1} > |a_n| \geq 0$ and $B \in O(A)$. If $t_k(B) = t_k(A)$ then

$$B = \begin{bmatrix} W & \\ & X_1 \end{bmatrix} A \begin{bmatrix} W^t & \\ & X_2 \end{bmatrix},$$

where $W \in \text{SO}_k$, $X_1, X_2 \in \text{SO}_{n-k}$.

Proof. Let $B = UAV$ where $U, V \in \text{SO}_n$ and write

$$U = (u_{ij}) = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}, \quad V = (v_{ij}) = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix},$$

where $U_{11}, V_{11} \in \mathbb{R}^{k \times k}$, $U_{22}, V_{22} \in \mathbb{R}^{(n-k) \times (n-k)}$. Denote

$$\begin{bmatrix} U_{11} & U_{12} \end{bmatrix} = \begin{bmatrix} u_{*1} & \dots & u_{*n} \end{bmatrix}, \quad \begin{bmatrix} V_{11} \\ V_{21} \end{bmatrix} = \begin{bmatrix} v_{1*} \\ \vdots \\ v_{n*} \end{bmatrix},$$

where $u_{*j}^t = (u_{1j}, \dots, u_{kj})$, $v_{j*} = (v_{j1}, \dots, v_{jk})$, $j = 1, \dots, n$. Then $t_k(UAV) = \text{tr}(U_{11}A_{11}V_{11} + U_{12}A_{22}V_{21}) = \text{tr}(A_{11}V_{11}U_{11} + A_{22}V_{21}U_{12}) = \sum_{i=1}^n a_i v_{i*} u_{*i}$. Since $v_{i*} u_{*i} \leq 1$, $\sum_{i=1}^n v_{i*} u_{*i} \leq k$ and $a_1 > \dots > a_k > \dots > a_n$, we have $\sum_{i=1}^n a_i v_{i*} u_{*i} \leq \sum_{i=1}^k a_{ii}$ with equality holds if and only if $v_{i*} u_{*i} = 1$ for $i \leq k$ and $v_{i*} u_{*i} = 0$ for $i > k$. Hence we have $v_{i*} = u_{*i}^t$ and $u_{*i} u_{*i}^t = 1$. Now $U = W \oplus X_1$ and $V = W^t \oplus X_1$ where $W \in \mathcal{O}_k$, $X_1, X_2 \in \mathcal{O}_{n-k}$ and $\det W = \det X_1 = \det X_2$. If $\det W = \det X_1 = \det X_2 = -1$, then we have $B = ((WD_1) \oplus (X_1 D_2)) A ((WD_1)^t \oplus (D_2 X_2))$ where $D_1 = I_{k-1} \oplus -1$ and $D_2 = -1 \oplus I_{n-k-1}$. \square

Thompson [9] gave the following result on characterizing the diagonal elements of $O(A)$.

Proposition 3.2. [9] *A vector $d = (d_1, \dots, d_n)$ is the diagonal of a matrix $A \in \mathbb{R}^{n \times n}$ with singular values $s_1 \geq s_2 \geq \dots \geq s_n$ if and only if d lies in the convex hull of those vectors $(\pm s_{\sigma(1)}, \dots, \pm s_{\sigma(n)})$ with an even number (possibly zero) of negative signs and arbitrary permutation σ .*

For matrices $A, B \in \mathbb{R}^{n \times n}$, the following result by Miranda and Thompson [8] can be regarded as a characterization of the extreme values of $O(A)$ under the linear map $X \mapsto \text{tr}(BX)$.

Proposition 3.3. [8] *Let $A, B \in \mathbb{R}^{n \times n}$ have singular values $s_1(A) \geq \dots \geq s_n(A)$ and $s_1(B) \geq \dots \geq s_n(B)$ respectively. Then*

$$\max_{U, V \in \text{SO}_n} \text{tr}(BUAV) = \sum_{i=1}^{n-1} s_i(A)s_i(B) + (\text{sign } \det(AB))s_n(A)s_n(B).$$

Theorem 3.4. *Let $A = \text{diag}(a_1, \dots, a_{n-1}, \pm a_n)$ where $a_1 > \dots > a_n \geq 0$ and $P = p_1 I_{n_1} \oplus \dots \oplus p_k I_{n_k}$ where $p_1 > \dots > p_k \geq 0$ and $n_1 + \dots + n_k = n$. Then*

(i) *if $p_k > 0$,*

$$\mathcal{G}_P(A) = \left\{ \begin{bmatrix} U_1 & & \\ & \ddots & \\ & & U_k \end{bmatrix} A \begin{bmatrix} U_1^t & & \\ & \ddots & \\ & & U_k^t \end{bmatrix} : \begin{array}{l} U_i \in \text{SO}_{n_i}, \\ i = 1, \dots, k \end{array} \right\};$$

(ii) *if $p_k = 0$,*

$$\mathcal{G}_P(A) = \left\{ \begin{bmatrix} U_1 & & \\ & \ddots & \\ & & U_{k-1} \\ & & & U \end{bmatrix} A \begin{bmatrix} U_1^t & & \\ & \ddots & \\ & & U_{k-1}^t \\ & & & V \end{bmatrix} : \begin{array}{l} U_i \in \text{SO}_{n_i}, \\ i = 1, \dots, k-1, \\ U, V \in \text{SO}_{n_k} \end{array} \right\}.$$

In both cases, $\mathcal{G}_P(A)$ is path connected.

Proof. (\supseteq) Obvious. (\subseteq). We assume that $A = A_1 \oplus \dots \oplus A_k$ where $A_i \in \mathbb{R}^{n_i \times n_i}$. We have $r(P, A) = d(P)^t d(A) = \sum_{i=1}^k p_i \text{tr} A_i$. Let $U, V \in \text{SO}_n$ such that $\text{tr}(PUAV) = r(P, A) = d(P)^t d(UAV)$. Write

$$UAV = B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1k} \\ B_{21} & B_{22} & \cdots & B_{2k} \\ \vdots & \cdots & \ddots & \vdots \\ B_{k1} & B_{k2} & \cdots & B_{kk} \end{bmatrix}$$

where $B_{ij} \in \mathbb{R}^{n_i \times n_j}$. We have $\text{tr}(PUAV) = \text{tr}(PB) = \sum_{i=1}^k p_i \text{tr} B_{ii}$. We shall show that $\text{tr} B_{ii} = \text{tr} A_i$ for all i whenever $p_i > 0$. By Proposition 3.2,

$d(B) = \sum \alpha_i s_i$ where $\alpha_i > 0$, $\sum \alpha_i = 1$ and s_i are vector of $(\pm a_{\sigma(1)}, \dots, \pm a_{\sigma(n)})$, σ is a permutation on $\{1, \dots, n\}$ and the number of negative signs is even (odd, respectively) if $\det A \geq 0$ (≤ 0 , respectively). If $k = 1$, then $P = p_1 I$, and the proof is trivial. Now consider $k > 1$, hence $p_1 > 0$. We first show that $\text{tr} B_{11} = \text{tr} A_1$. Note that $\text{tr} B_{11} < \text{tr} A_1$ holds if and only if at least one of the following cases hold:

- (1) there exists i_1 such that the first n_1 elements of s_{i_1} contain $-a_j$ where $j \leq n_1$;
- (2) there exists i_1 such that the first n_1 elements of s_{i_1} contain $\pm a_j$ where $j > n_1$.

In case (1), we construct s'_{i_1} from s by multiplying -1 to $-a_j$ and arbitrary a_q for some $q > n_1$. If in case (2), then there exists $i' < n_1$ such that $\pm a_{i'}$ will not be the first n_1 elements of s_{i_1} . In this case, we construct s'_{i_1} from s_{i_1} by interchanging $\pm a_j$ and $\pm a_{i'}$ and multiplying -1 to both if necessary to have $a_{i'}$ instead of $-a_{i'}$. Replace s_{i_1} in $\sum \alpha_i s_i$ by s'_{i_1} to form s . By Proposition 3.2, there exists $B' \in O(A)$ such that $d(B') = s$. We shall have $d(P)^t d(B) = d(P)^t (\sum \alpha_i s_i) = d(P)^t s + d(P)^t (s_{i_1} - s'_{i_1}) < d(P)^t s$, which contradicts the assumption on B . Therefore, we have $\text{tr} B_{11} = \text{tr} A_1$. By Lemma 3.1, we have $U = U_1 \oplus U_2$ and $V = V_1^t \oplus V_2$ where $U_1, V_1 \in \text{SO}_{n_1}$, $V_2, U_2 \in \text{SO}_{n-n_1}$ and $V_1^t = U_1$. Apply similar approach for B_{ii} where $p_i > 0$. Hence, if $p_k > 0$, we have $U = U_1 \oplus \dots \oplus U_k$ and $V = U^t$ where $U_i \in \text{SO}_{n_i}$, $i = 1, \dots, k$; otherwise if $p_k = 0$, $U = U_1 \oplus \dots \oplus U_{k-1} \oplus U'$ and $V = U_1^t \oplus \dots \oplus U_{k-1}^t \oplus V'$ where $U_i \in \text{SO}_{n_i}$, $i = 1, \dots, k-1$, $U', V' \in \text{SO}_{n_k}$. The path connectedness of $\mathcal{G}_P(A)$ follows from the path connectedness of SO_{n_i} for all i . \square

Corollary 3.5. *If $A \in \mathbb{R}^{n \times n}$ has n distinct singular values, then $L(O(A))$ has convex boundary for all linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$.*

Proof. Let $P, Q \in \mathbb{R}^{n \times n}$ be such that $\mathcal{L}(P, Q; O(A)) = L(O(A))$. Then $L(O(A))$ has convex boundary if for any $\theta \in [0, 2\pi]$, the set

$$\{-\sin \theta x + \cos \theta y : (x, y) \in \mathcal{L}(P, Q; O(A)), \cos \theta x + \sin \theta y = r_\theta\},$$

where $r_\theta = \max\{\cos \theta x + \sin \theta y : (x, y) \in \mathcal{L}(P, Q; O(A))\}$, is path connected. For any $\theta \in [0, 2\pi]$, we define $P'_\theta = -\sin \theta P + \cos \theta Q$ and $Q'_\theta = \cos \theta P + \sin \theta Q$, then we have

$$\begin{aligned} & \{-\sin \theta x + \cos \theta y : (x, y) \in \mathcal{L}(P, Q; O(A)), \cos \theta x + \sin \theta y = r_\theta\} \\ &= \{\text{tr}(P'_\theta U A V) : U, V \in \text{SO}_n, \text{tr}(Q'_\theta U A V) = r_\theta\} \\ &= \{\text{tr}(P'_\theta X) : X \in \mathcal{G}_{Q'_\theta}(A)\} \end{aligned}$$

Hence by Theorem 3.4, it is path connected. \square

Note that a set $M \subseteq \mathbb{R}^2$ is convex if and only if it is star-shaped and has convex boundary. Hence by Theorem 2.12 and Corollary 3.5, the following result is clear.

Theorem 3.6. *Let $n \geq 3$. If $A \in \mathbb{R}^{n \times n}$ has n distinct singular values, then $L(O(A))$ is convex for all linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$.*

In fact, the condition of distinct singular values in Theorem 3.6 can be removed by applying the following lemma.

Lemma 3.7. *Let $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^\ell$ be a linear map. Suppose $L(O(A))$ is convex for all A in a dense set S of $\mathbb{R}^{n \times n}$. Then $L(O(A))$ is convex for all $A \in \mathbb{R}^{n \times n}$.*

Proof. Suppose that $A_0 \in \mathbb{R}^{n \times n}$ such that $L(O(A_0))$ is not convex. Then there exist $x_1, x_2 \in L(O(A_0))$ such that $y = \frac{1}{2}(x_1 + x_2) \notin L(O(A_0))$. Since $L(O(A_0))$ is compact, there exists $\epsilon > 0$ such that $B(y, \epsilon) := \{x \in \mathbb{R}^\ell : \|x - y\| < \epsilon\}$ has empty intersection with $L(O(A_0))$. Since S is dense in $\mathbb{R}^{n \times n}$, there exists $A_\epsilon \in S$ such that for all $U, V \in \text{SO}_n$,

$$\|L(UA_0V) - L(UA_\epsilon V)\| < \frac{\epsilon}{2}.$$

Hence there exist $x'_1, x'_2 \in L(O(A_\epsilon))$ such that $\|x'_1 - x_1\| < \frac{\epsilon}{2}$ and $\|x'_2 - x_2\| < \frac{\epsilon}{2}$. By convexity of $L(O(A_\epsilon))$, $y' = \frac{1}{2}(x'_1 + x'_2) \in L(O(A_\epsilon))$. We have

$$\|y' - y\| = \left\| \frac{1}{2}(x'_1 + x'_2) - \frac{1}{2}(x_1 + x_2) \right\| < \frac{1}{2} \left(\frac{\epsilon}{2} + \frac{\epsilon}{2} \right) = \frac{\epsilon}{2}.$$

By assumption of A_ϵ , there exists $z \in L(O(A_0))$ such that $\|z - y'\| < \frac{\epsilon}{2}$. Then $\|z - y\| = \|(z - y') + (y' - y)\| < \|(z - y')\| + \|(y' - y)\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, contradicting the fact that $B(y, \epsilon) \cap L(O(A_0)) = \emptyset$. \square

Since the set of $n \times n$ matrices with n distinct singular values is dense in $\mathbb{R}^{n \times n}$, by Lemma 3.7 we have the following result.

Theorem 3.8. *Let $n \geq 3$. $L(O(A))$ is convex for all linear maps $L : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$ and $A \in \mathbb{R}^{n \times n}$.*

From the proof of Corollary 2.12, the convexity of $L(O(A))$ can be extended to $L(O_i(A_1, \dots, A_m))$, $i = 1, 2$.

Corollary 3.9. *Let $n \geq 3$. $L(O_i(A_1, \dots, A_m))$, $i = 1, 2$, is convex for all linear maps $L : (\mathbb{R}^{n \times n})^m \rightarrow \mathbb{R}^2$ and $A_1, \dots, A_m \in \mathbb{R}^{n \times n}$.*

References

- [1] W.S. Cheung, N.K. Tsing, The C -numerical Range of Matrices is Star-shaped, Linear and Multilinear Algebra 41 (1996), 245-250.
- [2] F. Hausdorff, Das Wertvorrat einer Bilinearform, Math. Zeit. 3 (1919), 314-316.
- [3] A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

- [4] C.K. Li, C -numerical Ranges and C -numerical Radii, Linear and Multilinear Algebra, 37 (1994), 51-82.
- [5] C.K. Li, Y.T. Poon, Convexity of the Joint Numerical Range, SIAM J. Matrix Analysis Appl. 21 (1999), 668-678.
- [6] C.K. Li, Y.T. Poon, Generalized Numerical Ranges and Quantum Error Correction, J. Operator Theory. 66 (2011), 335-351.
- [7] C.K. Li, T.Y. Tam, Numerical Ranges Arising from Simple Lie Algebras, Canad. J. Math. 52 (2000), 141-171.
- [8] H. Miranda , R.C. Thompson, Group Majorization, the Convex Hulls of Sets of Matrices, and the Diagonal Element-Singular Value Inequalities, Linear Algebra Appl. 199 (1994), 131-141.
- [9] R.C. Thompson, Singular Values, Diagonal Elements, and Convexity, SIAM J. Appl. Math. 32 (1977), 39-63.
- [10] O. Toeplitz, Das algebraische Analogon zu einem Satze von Fejer, Math. Zeit. 2 (1918), 187-197.
- [11] N.K. Tsing, On the Shape of the Generalized Numerical Ranges, Linear and Multilinear Algebra, 10 (1981), 173-182.
- [12] R. Westwick, A Theorem on Numerical Range, Linear and Multilinear Algebra, 2 (1975), 311-315